Optimising Magnetic Resonance Sampling Patterns for Parametric Characterisation

A. Reci, A.J. Sederman[a)], L.F. Gladden

Department of Chemical Engineering and Biotechnology, University of Cambridge, Philippa Fawcett Drive, Cambridge CB3 0AS, United Kingdom

[a)] Corresponding author

Name:	Andrew J. Sederman
Address:	Department of Chemical Engineering and Biotechnology
	University of Cambridge
	Philippa Fawcett Drive
	Cambridge CB3 0AS, University of Cambridge
Email:	ajs40@cam.ac.uk
Tel:	(+44) 1223 766338

[a)] Author to whom correspondence should be addressed. Electronic mail: ajs40@cam.ac.uk.

Abstract

Sampling strategies are often central to experimental design. Choosing efficiently which data to acquire can improve the estimation of parameters and reduce the acquisition time. This work is focused on designing optimal sampling patterns for Nuclear Magnetic Resonance (NMR) applications, illustrated with respect to the best estimate of the parameters characterising a lognormal distribution. Lognormal distributions are commonly used as fitting models for distributions of spin-lattice relaxation time constants, spin-spin relaxation time constants and diffusion coefficients. A method for optimising the choice of points to be sampled is presented which is based on the Cramér-Rao Lower Bound (CRLB) theory. The method's capabilities are demonstrated experimentally by applying it to the problem of estimating the emulsion droplet size distribution from a pulsed field gradient (PFG) NMR diffusion experiment. A difference of < 5% is observed between the predictions of CRLB theory and the PFG NMR experimental results. It is shown that CLRB theory is stable down to signal-to-noise ratios of ~10. A sensitivity analysis for the CRLB theory is also performed. The method of optimizing sampling patterns is easily adapted to distributions other than lognormal and to other aspects of experimental design; case studies of optimising the sampling scheme for a fixed acquisition time and determining the potential for reduction in acquisition time for a fixed parameter estimation accuracy are presented. The experimental acquisition time is typically reduced by a factor of 3 using the proposed method compared to a constant gradient increment approach that would usually be used.

Keywords: sampling pattern, lognormal distributions, Cramér-Rao Lower Bound theory, PFG NMR diffusion, emulsion droplet size distribution

1.Introduction

All experiments are to some degree limited by the amount of data that can be acquired. Choosing which acquired data will give the most statistically accurate parameters has long been a central topic in experimental design [1]. Recently, in the quest to achieve higher resolution (e.g. temporal resolution), mathematical techniques have been developed, of which the most prominent is Compressed Sensing [2, 3], which guarantee the accurate reconstruction of parameters from far fewer samples than was previously possible. As the number of samples acquired decreases significantly, the choice of which data are acquired becomes critical. The design of optimal sampling strategies depends on the application and it has been approached differently for example in Magnetic Resonance Imaging (MRI) [4], Nuclear Magnetic Resonance (NMR) spectroscopy [5], NMR relaxation time analysis [6, 7], electronic spectroscopy [8], X-ray ptychography [9] and Helium Atom Scattering [10].

This work is concerned with designing optimal sampling patterns for the most accurate estimate of parameters characterising a lognormal distribution. Lognormal distributions are ubiquitous in science and engineering, ranging from the description of the population distribution of organisms [11] to the size distribution of materials produced from particle processing techniques, such as in the food industry [12] and nanotechnology [13]. In NMR applications, lognormal distributions have been assumed to be a good approximation for polymer size distributions [14], molecular aggregate length distributions [15] and reverse micelle size distributions [16] obtained from pulsed field gradient (PFG) diffusion experiments; the spin-lattice relaxation time distribution of heavy oils obtained from fast-field cycling [17] and inversion recovery experiments [18]; and the spin-spin relaxation time distribution of heavy oils obtained from CPMG experiments [18, 19].

A systematic, statistical method is presented for designing sampling patterns when the distribution sought can be approximated by a lognormal distribution. The developed method is based on the Cramér-Rao Lower Bound (CRLB) theory [20]. The CRLB theory gives the theoretical minimum uncertainty in the estimation of the parameters of a model. This minimum uncertainty is the accuracy limit with which the parameters can be estimated, given the experimental data. The CRLB theory has previously been used to obtain optimal sampling patterns for mono-exponential decays in NMR relaxation time

analysis [21, 22], multidimensional COSY experiments [23] and diffusion-weighted MRI [24].

The proposed method is validated against PFG NMR diffusion experiments of an emulsion of toluene in water. The accurate measurement of the emulsion droplet size distribution is important in the food, pharmaceutical and oil recovery industry, among other areas [25]. Since its development [26], the measurement of the emulsion droplet size distribution using PFG NMR diffusion experiments has become an established characterisation technique [27-29]. The emulsion droplet size distribution obtained from PFG NMR diffusion experiments is commonly approximated to a lognormal distribution [26, 30-32]; this is supported by population balance statistics between droplet breakage and coalescence during emulsification [33] and by experimental results from other characterisation techniques such as dynamic light scattering [34] and confocal scanning laser microscopy [35].

The complete optimisation of an NMR experiment would also require the optimisation of specific NMR acquisition parameters and reconstruction techniques, which are separate from the optimisation of the sampling pattern considered in this work. The optimisation of specific NMR acquisition parameters and reconstruction techniques for PFG NMR diffusion experiments have been covered in detail elsewhere [24, 29, 36] and will not be addressed here.

Although the sampling method strategy presented here is illustrated with respect to improving the accuracy of estimation of lognormal distribution parameters, it can be easily adapted to other types of distributions, with the most obvious extension being distributions that can be approximated by the sum of lognormal distributions [16, 17, 29]. Of particular interest could be the optimization of sampling schemes for bi-exponential decays, the fitting of which is a long-standing challenge and remains a subject of debate [37-41]. In a related study [42], the application of CRLB theory to bi-exponential decays has been validated against experimental data.

The paper is structured as follows. Section 2 introduces the theory behind obtaining the emulsion droplet size distribution from PFG NMR diffusion experiments and the application of the CRLB theory to these experiments. The experimental sampling methods and PFG NMR setup are described in Section 3. The comparison between the predictions

of the CRLB theory and PFG NMR experimental results is presented in Section 4. The limitations, sensitivity and potential of the CRLB theory are also discussed in Section 4.

2.Theory

The structure of this Section is as follows. Section 2.1 briefly reviews the theory behind the extraction of the emulsion droplet size distribution from PFG NMR diffusion experiments. Section 2.2 introduces the CRLB theory in its generality, and in Section 2.3 the CRLB theory is applied to the problem of optimizing the sampling pattern for the PFG NMR diffusion acquisitions of emulsion droplet size distributions.

2.1. PFG NMR diffusion of emulsion systems

The ideal NMR signal attenuation acquired from a PFG NMR diffusion experiment of an unconstrained component, for a range of pulsed field gradient amplitudes, g_i (i = 1, 2, ..., n), is described by the Stejskal-Tanner equation [43]:

$$y_i = A \exp\left(-\gamma^2 {g_i}^2 \delta^2 D \left(\Delta - \frac{\delta}{3}\right)\right), \quad (1)$$

where γ is the gyromagnetic ratio of the NMR-active nucleus, δ is the pulsed field gradient duration, Δ is the diffusion time, D is the unconstrained diffusion coefficient and A is a scaling factor.

If the component is constrained in the dispersed phase of an emulsion and the diffusion time is such that the root mean square distance travelled due to Brownian motion is larger than the characteristic size of the droplet, the apparent diffusion coefficient is smaller than the unconstrained diffusion coefficient, with the apparent value depending on the droplet size. Since the emulsion is characterized by a distribution of droplet sizes, the PFG NMR signal attenuation is a multi-exponential decay, from which the droplet size distribution can be extracted. The ideal NMR signal attenuation acquired from the dispersed phase of an emulsion with droplet number size distribution $f_0(a_j)$, has been calculated by Murday and Cotts [44]:

$$y_i = A \sum_{j=1}^{p} {a_j}^3 f_0(a_j)\, R(g_i, a_j), \quad (2a)$$

where a_j (j = 1, 2, ..., p) is the discretized list of droplet radii and A is a scaling factor. The standard notation for the number size distribution, $f_0(a_j)$, [45] has been used. The factor a_j^3 could be combined with $f_0(a_j)$ to give the volume size distribution, $f_3(a_j)$, but we choose to focus on the number size distribution, for ease of interpretation. In Eq. (2a):

$$R(g_i, a_j) = \exp\left(-2\gamma^2 {g_i}^2 \sum_{k=1}^{\infty} \frac{1}{\alpha_k^2\left(\alpha_k^2 a_j^2 - 2\right)} r(\alpha_k)\right), \tag{2b}$$

where:

$$r(\alpha_k) = \frac{2\delta}{\alpha_k^2 D} + \frac{-2 - \exp\left(-\alpha_k^2 D(\Delta-\delta)\right) + 2\exp\left(-\alpha_k^2 D\Delta\right) + 2\exp\left(-\alpha_k^2 D\delta\right) - \exp\left(-\alpha_k^2 D(\Delta+\delta)\right)}{\left(\alpha_k^2 D\right)^2}. \tag{2c}$$

In Eqs. 2(b, c), D is the unconstrained diffusion coefficient of the dispersed phase and α_k are the solutions to the equation:

$$J_{3/2}(\alpha_k a_j) = \alpha_k a_j J_{5/2}(\alpha_k a_j), \tag{2d}$$

where J_k is the Bessel function of the first kind and of order k. As discussed in Section 1, the droplet number size distribution, $f_0(a_j)$, is typically well approximated to a lognormal distribution:

$$f_0(a_j) = \frac{1}{a_j s_g \sqrt{2\pi}} \exp\left(-\frac{\left(\ln a_j - \bar{a}_{0,0}\right)^2}{2{s_g}^2}\right), \tag{3}$$

where $\bar{a}_{0,0}$ is the geometric mean of the droplet size distribution and s_g is the geometric standard deviation of the droplet size distribution, following standard notation (the arithmetic mean of the droplet size distribution is $\bar{a}_{1,0}$ and the standard deviation of the droplet size distribution is s). If the droplet number size distribution is lognormal, the volume size distribution is also lognormal [45, 46]. As a result, a similar analysis of what follows in the next sections can also be applied to the volume size distribution, $f_3(a_j)$.

The ideal NMR signal, therefore, depends on three parameters: the scaling factor, A, the geometric mean, $\bar{a}_{0,0}$, and the geometric standard deviation, s_g, of the lognormal droplet size distribution. The accurate estimation of these parameters is the objective of the PFG NMR diffusion experiment.

2.2. CRLB theory

All signals are to some degree corrupted by noise. As a result, the acquired signal, $\hat{y}_i$, is composed of the ideal signal, y_i, and an unknown noise term, ϵ_i, according to:

$$\hat{y}_i = y_i + \epsilon_i \ . \tag{4}$$

In this work, it will be assumed that ϵ_i are random and independently normally distributed with zero mean and variance σ^2 (Gaussian noise). Suppose the ideal signal, y_i, depends on q parameters, θ_l (l = 1, 2, ..., q). In the presence of noise, the estimate of parameters, θ_l, becomes uncertain and there is an uncertainty associated with the estimate. A common characteristic of this uncertainty is the standard deviation, std(θ_l). Accurate estimates of θ_l have a small std(θ_l). The lower limit of this standard deviation is given by the Cramér-Rao Lower Bound (CRLB) theory:

$$\mathrm{std}(\theta_k) \geq \sqrt{\left(\underline{F}^{-1}\right)_{k,k}} \ , \tag{5}$$

where $\underline{F}$ is the Fisher Information Matrix [47], whose content summarizes the amount of information available in the acquired signal about the estimation of the parameters. Although the CRLB theory refers to the lower limit in the uncertainty associated with a given parameter, this lower limit is practically the actual uncertainty associated with that parameter, as long as the method used to estimate that parameter is unbiased (the CRLB theory needs to be altered in order to be applied to biased methods of estimation, such as regularization). The Fisher Information Matrix is related to the likelihood function, L, by:

$$F_{l_1,l_2} = \mathrm{E}\left(\frac{\partial L}{\partial \theta_{l_1}} \frac{\partial L}{\partial \theta_{l_2}}\right) , \tag{6}$$

where E refers to the expected value, and the likelihood function related to the random variables, ϵ_i, is defined as:

$$L = \ln\left(\prod_{i=1}^{n} P(\epsilon = \epsilon_i)\right) = n \ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \ . \tag{7}$$

When the likelihood function in Eq. (7) is applied to the definition of the Fisher Information Matrix in Eq. (6), a simplified expression for the Fisher Information Matrix is obtained [48]:

$$F_{l_1,l_2} = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n}\left(\frac{\partial y_i}{\partial \theta_{l_1}} \frac{\partial y_i}{\partial \theta_{l_2}}\right)\right) . \tag{8}$$

The combination of Eqs. (5) and (8) can be used to estimate the uncertainty associated with the estimation of each of the model parameters. The co-factor matrix properties can be further used to simplify the calculations [49].

2.3. Application of CRLB theory to emulsion systems

Combining Eqs. (2a) and (3), the ideal PFG NMR diffusion signal, y_i, is given by:

$$y_i = A\sum_{j=1}^{p}\frac{{a_j}^2}{s_g\sqrt{2\pi}}\exp\left(-\frac{\left(\ln a_j-\bar{a}_{0,0}\right)^2}{2{s_g}^2}\right)R(g_i,a_j), \tag{9}$$

and, therefore, depends on three parameters: A, $\bar{a}_{0,0}$ and s_g. In the presence of noise, the uncertainty associated with the estimate of each of these parameters can be calculated from the CRLB theory, according to Eqs. (5) and (8). To apply Eq. (8), the partial derivatives of y_i with respect to A, $\bar{a}_{0,0}$ and s_g are needed, which can be calculated from Eq. (9):

$$\frac{\partial y_i}{\partial A} = \sum_{j=1}^{p}\frac{{a_j}^2}{s_g\sqrt{2\pi}}\exp\left(-\frac{\left(\ln a_j-\bar{a}_{0,0}\right)^2}{2{s_g}^2}\right)R(g_i,a_j)\ , \tag{10a}$$

$$\frac{\partial y_i}{\partial \bar{a}_{0,0}} = A\sum_{j=1}^{p}\frac{{a_j}^2\left(\ln a_j-\bar{a}_{0,0}\right)}{{s_g}^3\sqrt{2\pi}}\exp\left(-\frac{\left(\ln a_j-\bar{a}_{0,0}\right)^2}{2{s_g}^2}\right)R(g_i,a_j)\ , \tag{10b}$$

$$\frac{\partial y_i}{\partial s_g} = A\sum_{j=1}^{p}\frac{{a_j}^2\left[\left(\ln a_j-\bar{a}_{0,0}\right)^2-{s_g}^2\right]}{{s_g}^4\sqrt{2\pi}}\exp\left(-\frac{\left(\ln a_j-\bar{a}_{0,0}\right)^2}{2{s_g}^2}\right)R(g_i,a_j)\ . \tag{10c}$$

The important result is that the uncertainty associated with the estimate of each parameter depends on A, $\bar{a}_{0,0}$, s_g, σ and most importantly on the sampling pattern, g_i. As a result, given the noise level and an experimental system (characterised by A, $\bar{a}_{0,0}$ and s_g), the error associated with the estimate of each parameter depends solely on the sampling pattern, g_i. Therefore, by optimizing the sampling pattern, it is possible to decrease to a minimum the uncertainty associated with the estimate of a given parameter, hence, making the estimate of that parameter as accurate as possible.

The optimal sampling pattern for the most accurate estimation of one of the parameters is not necessarily the optimal sampling pattern for the most accurate estimation of another parameter. Therefore, it is necessary to know what the most useful information for the particular application is, and to express it in terms of an objective function. This objective function could simply be the minimization of std(A), std($\bar{a}_{0,0}$) or std(s_g), or more complicated expressions such as different moments of the lognormal distribution [50]. The

present method can cope with different objective functions. In this work, the choice is made to minimize the maximum error bar associated with each individual point in the droplet size number distribution, given by:

$$\chi = \max_j \left[\mathrm{std}\left(f_0(a_j)\right)\right] \approx \max_j \left[\left|\frac{\partial f_0(a_j)}{\partial \bar{a}_{0,0}}\mathrm{std}(\bar{a}_{0,0})\right| + \left|\frac{\partial f_0(a_j)}{\partial s_\mathrm{g}}\mathrm{std}(s_\mathrm{g})\right|\right]. \quad (11)$$

The derivation of Eq. (11) is outlined in the Appendix. Using this definition, the best sampling pattern has the minimum objective function, χ. The choice has been made purely to demonstrate the capabilities of the method, but other objective functions can be defined. It can be shown that $\mathrm{std}(\bar{a}_{0,0}) \propto \sigma A^{-1}$, $\mathrm{std}(s_\mathrm{g}) \propto \sigma A^{-1}$ and therefore $\chi \propto \sigma A^{-1}$. The ratio A/σ determines the signal-to-noise ratio. As a result, for a given signal-to-noise ratio, the objective function, χ, depends solely on the lognormal distribution parameters $\bar{a}_{0,0}$ and s_g and the sampling pattern, g_i. In the following discussion, when stating σ, the work is referring to the noise level for a signal normalized to a maximum of 1.

3.Materials and methods

3.1. Experimental

An emulsion of toluene dispersed in water was prepared as follows. A non-ionic surfactant, Triton^{TM} X-100 (Sigma-Aldrich, laboratory grade), was first dissolved in deionised water. Toluene (Sigma-Aldrich, ≥ 99 % purity) was then mixed into the solution using a Janke & Kunkel RW-20 stirrer, equipped with a 2 cm radius stainless steel blade, driven at 150 rpm for 10 min. The emulsion formed was composed of 48.5 wt% water, 48.5 wt% toluene and 3 wt% surfactant. A small volume from the batch was transferred into a 5 mm NMR tube and a PFG NMR diffusion experiment was performed on this sample. The measurement time was 2.5 h. A second acquisition immediately after the first produced the same emulsion droplet size distribution, thereby confirming that the emulsion was stable during the course of the experiment.

Experiments were performed at 20 ± 0.5 °C on a Bruker DMX 300 spectrometer, operating at a resonant frequency of 300.13 MHz for ^{1}H nucleus observation. The maximum gradient amplitude available was 1176 G cm^{-1} and the radiofrequency (RF) coil had a diameter of 5 mm.

The unconstrained diffusion coefficient of toluene, D, was measured to be $(2.06 \pm 0.02) \times 10^{-9}$ m^2 s^{-1} using a stimulated echo PFG NMR pulse sequence [51], with pulsed field gradient duration δ = 1.5 ms, diffusion time Δ = 200 ms and the pulsed field gradient amplitude, g_i, in the range 0.1 – 40 G cm^{-1} (16 equidistant points). This is in agreement with reported values at this temperature [52].

The same pulse sequence was then applied to the emulsion system, with the following acquisition parameters: δ = 5 ms, Δ = 400 ms, g_i = 4.8 – 300 G cm^{-1} (252 equidistant points). The attenuation of the aromatic peak intensity only was followed to estimate the droplet size distribution. The contribution to the aromatic peak intensity from the surfactant, as compared to toluene, is small (calculated to be 0.9%) and was taken into account by subtracting it from the overall aromatic peak intensity at all the gradient values used. The emulsion droplet size distribution was reconstructed for a discrete list of droplet radii, a, composed of 32 points equidistantly spaced between 0.1-15 μm, as was used in the work of Hollingsworth and Johns [29]. Only the first 6 terms were kept from the solutions of the Bessel equation in Eq. 2(d); it was observed that keeping more terms made no difference to the droplet size distribution predicted.

3.2. Sampling patterns

A sampling pattern for the acquisition of the PFG NMR data consists of the choice of g_i, the list of pulsed field gradients at which the attenuation of the NMR signal is acquired. This involves choosing the number of pulsed field gradients used, n, and their respective values. n is typically a user-defined variable and larger values of n will give more accurate estimate of parameters. In this work, n is kept fixed at 32 except when combining n and σ in a way which keeps either the experimental acquisition time or the parameter estimation accuracy fixed; these results are reported in sections 4.3 and 4.4 respectively.

For a given number of points, n, the optimal choice of g_i would require setting the partial derivatives of the expression in Eq. (11) with respect to each individual g_i to zero and solving the set of resulting n simultaneous non-linear equations; this is computationally intractable. As a result, this paper focuses on finding the optimal sampling pattern among classes of practical sampling patterns. Two classes of sampling patterns are considered: power law spacing of the linear version of sampled points and power law spacing of the logarithmic version of sampled points.

The power law spacing of the linear version of sampled points, which for simplicity will be referred to as a linear sampling pattern, takes the form:

$$g_i = g_1 + (g_n - g_1)\left(\frac{i-1}{n-1}\right)^r, \tag{12}$$

where g_1 and g_n are the first and last pulsed field gradients used and r determines the density distribution of the data points; $r > 1$ has a higher density of points at the start of the decay, $r < 1$ has a higher density of points at the end of the decay and $r = 1$ refers to equidistant pulsed field gradient values. Acquiring the point that has the largest amplitude is always beneficial because it contains the most information about the attenuation of the signal [21, 22, 53]. Therefore, g_1 was kept at the minimum pulsed field gradient used experimentally of 4.8 G cm^{-1}. As a result, optimization of the linear sampling schemes effectively consists of optimizing g_n and r.

The power law spacing of the logarithmic version of sampled points, which for simplicity will be referred to as a logarithmic sampling pattern, takes the form:

$$\log_{10} g_i = \log_{10} g_1 + (\log_{10} g_n - \log_{10} g_1)\left(\frac{i-1}{n-1}\right)^r, \tag{13}$$

where the parameters are defined similarly to the linear sampling pattern. Using the same assumptions as for the linear sampling scheme, optimization of the logarithmic sampling schemes consists of optimizing g_n and r.

The comparison between the predictions of the CRLB theory and experimental results is performed as follows. Using a range of g_n from 20 to 300 G cm^{-1} and a range of r from 0.1 to 10, 10^4 linear sampling patterns and 10^4 logarithmic sampling patterns with $n = 32$ points were simulated using Eq. (12) and Eq. (13), respectively. For each of these sampling patterns, the CRLB theory was used to predict the objective function, χ, defined in Eq. (11). This was compared to the objective function obtained from the experimental data which was calculated as follows. For each of the simulated sampling patterns, the closest experimental 32 g_i points (out of the total 252 g_i points acquired) were selected. The experimental attenuation data corresponding to these 32 points was then fitted to a lognormal distribution for the droplet size distribution. std($\bar{a}_{0,0}$) and std(s_g) were extracted from the fitting procedure and used to calculate the experimental objective function, χ.

The difference between the simulated and the experimental sampling patterns was negligible. This ensures that the comparison between the predictions of the CRLB theory

(which is performed using the simulated sampling pattern) and the experimental results (which uses the selected experimental points) is unaffected by mismatches in the corresponding sampling patterns used.

4.Experimental results and discussion

Initially, the full experimental results for the PFG NMR diffusion experiment of the emulsion of toluene in water are given. Then, the effect of choosing different sampling patterns in the estimation of the lognormal distribution parameters of the emulsion droplet size distribution is investigated by comparing the predictions of the CRLB theory with experimental data. The limitations and sensitivity of the CRLB theory are subsequently discussed. Lastly, the CRLB theory is used to optimise the sampling pattern for a fixed acquisition time experiment and to determine the potential for reduction in acquisition time, while keeping the parameter estimation accuracy fixed.

4.1. Validation of CRLB theory from experimental data

The signal attenuation from the aromatic peak of toluene in the PFG NMR diffusion experiment of the emulsion of toluene in water is shown in Fig. 1(a). A curvature is observed in the plot of the signal against the sampling pattern, g_i^2. This confirms that toluene is experiencing restricted diffusion resulting from being confined within droplets. The signal attenuation is fitted to a lognormal distribution for the droplet size distribution; the best fit lognormal distribution with $\bar{a}_{0,0} = 1.09$, $s_g = 0.38$ ($\bar{a}_{1,0} = 3.2$ μm, $s = 1.2$ μm) is shown in Fig. 1(b). The error associated with each individual point in the discretized lognormal distribution is negligible because of the relatively small σ (estimated to be 5.6×10^{-4}) and the large number of points in the decay (252).

From the 252 points in the decay, sampling schemes with 32 points are selected systematically according to linear sampling patterns in Eq. (12), using the method described in section 3.2. The objective function, χ, defined in Eq. (11), is calculated for each of these sampling schemes and is compared to the predictions of the CRLB theory for the objective function. A small objective function indicates a good sampling pattern. The results are shown in Fig. 2. Fig. 2(a) shows the contour map of χ as a function of linear

sampling pattern parameters g_n and r, as predicted by the CRLB theory. Fig. 2(b) shows a similar contour map obtained from the experimental data. There is a very good agreement between the CRLB and experimental contour maps; on average there is < 5% difference between the corresponding χ values of the two maps, thereby validating the application of the CRLB theory to the optimization of sampling patterns. For a given value of r, for example $r = 1$ (corresponding to equidistant points), the value of the objective function is large at small g_n because late decay points are not sampled and, therefore, the uncertainty related to estimating the small droplet sizes (corresponding to small apparent diffusion coefficients) is large. The value of the objective function is also large at very large g_n because early decay points are not sampled properly and, therefore, the uncertainty related to estimating the large droplet sizes (corresponding to large apparent diffusion coefficients) is large. As a result, for any given r, there is an optimum g_n which minimises the objective function. Similar arguments can be made for a fixed g_n and variable r. The best linear sampling pattern for 32 sampled points is predicted by the CRLB theory at $r = 0.67$ and $g_n = 159$ G cm^{-1}, at which point $\chi = 1.03 \times 10^{-3}$.

A similar comparison is made between the predictions of the CRLB theory and the experimental data for logarithmic sampling patterns of 32 points selected according to Eq. (13). The results are shown in Fig. 3. Again, very good agreement is observed between the contour maps; on average there is < 5% difference between the corresponding χ values of the two maps, which further validates the application of the CRLB theory to the optimization of sampling patterns. The best logarithmic sampling pattern for 32 sampled points is predicted by the CRLB theory at $r = 0.23$ and $g_n = 161$ G cm^{-1}, at which point $\chi = 1.03 \times 10^{-3}$. It is noted that the minimum objective function value for linear sampling patterns is equal to the minimum objective function value for logarithmic sampling patterns ($\chi = 1.03 \times 10^{-3}$), although the two classes of sampling patterns are very different. This is an indication that this minimum objective function is a global minimum (with respect to all possible classes of sampling patterns). Given this observation, in the following sections, the discussion will be limited to linear sampling patterns. For completeness it is noted that, as discussed in Section 3.2, the proof that this minimum objective function is a global minimum for all sampling patterns would require the solution of an intractable computational problem.

4.2. Limitations and sensitivity

Discrepancies between the CRLB theory predictions and experimental data have been previously reported at large σ (low signal-to-noise ratios) [21]. Considering the increasing interest in low-field measurements, where the signal-to-noise ratio may limit the extraction of parameters from the NMR signal, it is important to find the limit of applicability of the CRLB theory with respect to the signal-to-noise ratio. The following discussion aims to find the limit of applicability of the CRLB theory to the present problem of estimating the lognormal distribution parameters of the emulsion droplet size distribution.

Monte Carlo type simulations [54] were performed as follows. The 32 experimental data points acquired at the optimal linear sampling pattern, reported in Section 4.1, were extracted. Random Gaussian noise was added to these acquired points, such that noise levels, σ, in the range 10^{-3}-5×10^{-1} were studied. For each noise level, 100 signals were simulated, differing only in the randomness of noise. Each of these signals was fitted to a lognormal distribution for the emulsion droplet size distribution. The values of $\bar{a}_{0,0}$ and s_g from each of these fittings were then used to estimate std($\bar{a}_{0,0}$) and std(s_g). These standard deviations were then used to calculate the objective function, χ, according to Eq. (11). This was then compared to the CRLB predictions. The results are shown in Fig. 4(a). It is observed that there is very good agreement between the Monte Carlo simulations and the CRLB theory predictions up to $\sigma \sim 0.1$ (signal-to-noise ratio ~10). This limit will be slightly different for systems characterised by different $\bar{a}_{0,0}$ and s_g. However, a general conclusion can be made that the method is stable even at highly noisy data (i.e. signal-to-noise ratios of ~10).

As discussed in Section 2, the application of the CRLB theory requires an initial guess of the lognormal distribution parameters, $\bar{a}_{0,0}$ and s_g. A valid question is: How accurate does this guess need to be? Since it is easier to work in terms of the arithmetic mean, $\bar{a}_{1,0}$ (μm), and standard deviation, s (μm), it is more sensible to ask about the accuracy with which $\bar{a}_{1,0}$ and s need to be *a priori* known. To answer this question, a sensitivity analysis was performed as follows. Assume that the optimal linear sampling pattern of 32 points found in Section 4.1 was designed on the basis that $\bar{a}_{1,0}$ = 3.2 μm, s = 1.2 μm. Now, assume that this estimate for $\bar{a}_{1,0}$ and s was wrong, and the true value for the arithmetic mean and standard deviation of the lognormal distribution was actually $\bar{a}_{1,0t}$ and s_t. For these true values of $\bar{a}_{1,0t}$ and s_t, the CRLB theory can be used to find the minimum objective function,

χ_t, corresponding to the ideal linear sampling pattern, and the actual value of χ evaluated for the actual sampling pattern used. The percentage difference of χ from χ_t is an indication of how sensitive the CRLB theory is to wrong initial guesses of the lognormal distribution parameters. A contour map of the percentage difference of χ from χ_t, as a function of the percentage difference of $\bar{a}_{1,0}$ from $\bar{a}_{1,0t}$ and s from s_t is shown in Fig. 4(b). The key observation is that the CRLB theory is relatively insensitive to the guess of the standard deviation, s, as compared to the sensitivity to the arithmetic mean, $\bar{a}_{1,0}$. This is important as in general it is more difficult to guess s than $\bar{a}_{1,0}$. With regards to the sensitivity to the guess of $\bar{a}_{1,0}$, the CRLB theory is relatively insensitive around the true value, $\bar{a}_{1,0t}$, but becomes increasingly more sensitive the further away from the true value the guess is.

The versatility of the CRLB theory to be adapted to different objectives is now considered with respect to two cases studies; in Section 4.3 the sampling scheme is optimised while keeping the acquisition time fixed, while in Section 4.4 the potential for reduction in acquisition time is investigated while keeping the parameter estimation accuracy fixed.

4.3. Optimising the sampling scheme for a fixed acquisition time

Optimization of sampling patterns has so far been made for a fixed number of points, $n = 32$. As discussed in section 3.2, for a fixed noise level σ, acquiring more points will lead to a more accurate estimate for the lognormal distribution parameters. However, this comes at the cost of a longer acquisition time. If the acquisition time is limited by, for example, the stability of the emulsion under investigation, is it better to acquire fewer points with low σ (high signal-to-noise ratio), or more points with higher σ (low signal-to-noise ratio)?

The CRLB theory can be used to answer this question as follows. In a PFG NMR diffusion experiment, the acquisition time, t, is related to the number of points acquired, n, and the noise level, σ, by $t \propto n\ \sigma^{-2}$ [55]. Therefore, for a fixed acquisition time, $\sigma \propto n^{1/2}$. For a given n and the corresponding σ (for a fixed acquisition time), the CRLB theory can be used to determine the minimum objective function, χ, for linear sampling patterns, as was done to produce Fig. 2(a). This is repeated for n in the range 3-32, and the plot of the minimum χ, as a function of n is shown in Fig. 5(a). At large n, the acquired data are noisy and therefore the best sampling pattern for that n has a relatively large minimum χ. At small n, although the data are of a high signal-to-noise ratio, there are not many points to

fit the lognormal distribution parameters to. As a result, the minimum χ is again relatively large. The optimum linear sampling pattern is to acquire 7 points with linear sampling pattern parameters r = 0.98 (close to being equidistant points) and g_n = 147 G cm^{-1}, which is illustrated in Fig. 5(b).

4.4 Reducing the acquisition time for a fixed parameter estimation accuracy

While the parameter estimation accuracy achieved using a default sampling pattern may be acceptable in an experiment, it is important to know whether the same accuracy can be achieved for a shorter acquisition time. This problem will be investigated with respect to a hypothetical experiment, but the idea extends easily to other experiments. It is assumed that the system under the investigation is the emulsion used in this paper, and that the default sampling pattern that was initially being used to study the emulsion by a PFG NMR diffusion experiment was composed of 16 points equidistant from each other (linear sampling pattern with r = 1), with the largest gradient value being g_n = 300 G cm^{-1}. Using the CRLB theory, the objective function, χ_{def}, that characterizes the parameter estimation accuracy of this default sampling pattern can be calculated after a reference noise level, σ_{ref}, is defined. As discussed in Section 4.3, the acquisition time for this sampling pattern is given by $t_{\text{def}} \propto 16\ \sigma_{\text{ref}}^{-2}$.

The reduction ratio in acquisition time, t / t_{def}, which can be achieved by better sampling, while keeping $\chi = \chi_{\text{def}}$ is calculated as follows. For a given n, the CRLB theory can be used to determine the minimum objective function, χ, as a function of the noise level, σ, for linear sampling patterns, as was done to produce Fig. 2(a). The noise level is then tuned such that $\chi = \chi_{\text{def}}$. Using the tuned noise level, the acquisition time for the chosen sampling pattern is calculated as $t \propto n\ \sigma^{-2}$. The process is repeated for n in the range 3-32 and the variation of t / t_{def} with n is shown in Fig. 5(a). It is observed that the minimum acquisition time is achieved for a linear sampling pattern of n = 7 points with linear sampling pattern parameters r = 0.98 (close to being equidistant points) and g_n = 147 G cm^{-1}, which is illustrated in Fig. 5(b). The reduction ratio in acquisition time achieved was t / t_{def} = 0.35, corresponding to an acceleration factor of ~3.

The fact that the same optimal sampling pattern is achieved for both the process of keeping t fixed and minimizing χ (discussed in Section 4.3) and the process of keeping χ fixed and minimizing t is no coincidence. Indeed, it can be easily shown that $t_{\text{fixed }\chi} \propto \chi^2_{\text{fixed }t}$, which

guarantees the same optimum *n* value where the minimum of either of these two functions is achieved.

5. Conclusions

A method was presented, based on the Cramér-Rao Lower Bound theory, for designing optimal sampling patterns for NMR experiments, illustrated with respect to the most accurate estimation of lognormal distribution parameters. The method was validated against experimental data of a PFG NMR diffusion experiment of an emulsion of toluene in water, used to obtain the emulsion droplet size distribution. The difference between the predictions of the CRLB theory and experimental data was < 5% and it was shown that the CRLB theory is stable even at signal-to-noise ratios of ~10. Using the CRLB theory, it was demonstrated that for a PFG NMR experiment which usually uses constant spacing between gradient increments, the acquisition time can be reduced by a typical factor of 3 while keeping the parameter estimation accuracy unchanged, through optimization of the sampling pattern. The presented approach can guide experimental design in wide range of experiments.

Acknowledgements

A. R. acknowledges Gates Trust Cambridge for financial support.

Appendix: Derivation of Eq. (11)

Let *x*, *z* be two random variables, which may be correlated, and *h*(*x*, *z*) be a function of these two random variables. Then, from the truncated Taylor series expansion [56]:

$$\mathrm{var}(h) \approx \left(\frac{\partial h}{\partial x}\right)^2 \mathrm{var}(x) + \left(\frac{\partial h}{\partial z}\right)^2 \mathrm{var}(z) + 2\frac{\partial h}{\partial x}\frac{\partial h}{\partial z}\mathrm{cov}(x,z), \quad \text{(A1)}$$

where var refers to the variance of a random variable and cov refers to the covariance between two random variables. Since the correlation coefficient between two random variables is always ≤ 1, then:

$$\mathrm{cov}(x,z) \leq \mathrm{std}(x)\,\mathrm{std}(z). \tag{A2}$$

From the relationship between the variance and standard deviation, Eq. (A1) can then be written as:

$$\left(\mathrm{std}(h)\right)^2 \leq \left(\frac{\partial h}{\partial x}\right)^2 \left(\mathrm{std}(x)\right)^2 + \left(\frac{\partial h}{\partial z}\right)^2 \left(\mathrm{std}(z)\right)^2 + 2\left|\frac{\partial h}{\partial x}\right|\left|\frac{\partial h}{\partial z}\right| \mathrm{std}(x)\mathrm{std}(z) = \left(\left|\frac{\partial h}{\partial x}\right| \mathrm{std}(x) + \left|\frac{\partial h}{\partial z}\right| \mathrm{std}(z)\right)^2. \tag{A3}$$

It follows that a reasonable approximation for the lower limit of std(h) is:

$$\mathrm{std}(h) \approx \left|\frac{\partial h}{\partial x}\right| \mathrm{std}(x) + \left|\frac{\partial h}{\partial z}\right| \mathrm{std}(z). \tag{A4}$$

References

[1] R. A. Fisher, Statistical Methods for Research Workers, Oliver and Boyd, 10th ed., 1948.

[2] E. J. Candés, J. Romberg and T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, IEEE Trans. Inf. Theory 52 (2006), 489-509.

[3] D. L. Donoho, Compressed Sensing, IEEE Trans. Inf. Theory, 52 (2006), 1289-1306.

[4] M. Lustig, D. Donoho and J. M. Pauly, Sparse MRI: The application of Compressed Sensing for rapid MR imaging, Magn. Reson. Med. 58 (2007), 1182-1195.

[5] M. Mobli, M. W. Maciejewski, A. D. Schuyler, A. S. Stern and J. C. Hoch, Sparse sampling methods in multidimensional NMR, Phys. Chem. Chem. Phys. 14 (2012), 10835-10843.

[6] C. K. Anand, A. D. Bain, A. Sharma, Optimized sampling patterns for multidimensional T_2 experiments, J. Magn. Reson. 197 (2009), 63-70.

[7] Y.-Q. Song, Y. Tang, M. D. Hürlimann and D. G. Cory, Real-time optimization of nuclear magnetic resonance experiments, J. Magn. Reson. 289 (2018), 72-78.

[8] S. Roeding, N. Klimovich and T. Brixner, Optimizing sparse sampling for 2D electronic spectroscopy, J. Chem. Phys. 146 (2017), 084201.

[9] T. B. Edo, D. J. Batey, A. M. Maiden, C. Rau, U. Wagner, Z. D. Pešić, T. A. Waigh and J. M. Rodenburg, Sampling in X-ray ptychography, Phys. Rev. A 87 (2013), 053850.

[10] A. Jones, A. Tamtögl, I. Calvo-Almazán and A. Hansen, Continuous Compressed Sensing for surface dynamical processes with Helium Atom Scattering, Sci. Rep. 6 (2016), 27776.

[11] A. L. Koch, The logarithm in biology. 1. Mechanisms generating the log-normal distribution exactly, J. Theoret. Biol. 12 (1966), 276-290.

[12] E. Limpert, W. A. Stahel and M. Abbt, Log-normal distributions across the sciences: Keys and clues, BioScience 51 (2001), 341-352.

[13] L. B. Kiss, J. Söderlund, G. A. Niklasson and C. G. Granqvist, New approach to the origin of lognormal size distributions of nanoparticles, Nanotechnology 10 (1999), 25-28.

[14] N. H. Williamson, M. Nydén and M. Röding, The lognormal distribution models for estimating molecular weight distributions of polymers using PGSE NMR, J. Magn. Reson. 267 (2016), 54-62.

[15] A. J. Baldwin, S. J. Anthony-Cahill, T. P. J. Knowles, G. Lippens, J. Christodoulou, P. D. Barker and C. M. Dobson, Measurement of myloid fibril length distributions by inclusion of rotational motion in solution NMR diffusion measurements, Agnew. Chem. Int. Ed. 47 (2008), 3385-3387.

[16] S. J. Law and M. M. Britton, Sizing of reverse micelles in microemulsions using NMR measurements of diffusion, Langmuir 28 (2012), 11699-11706.

[17] J.-P. Korb, A. Louis-Joseph and L. Benamsili, Probing structure and dynamics of bulk and confined crude oils by multiscale NMR spectroscopy, diffusometry, and relaxometry, J. Chem. Phys. B 117 (2013), 7002-7014.

[18] L. Benamsili, J.-P. Korb, G. Hamon, A. Louis-Joseph, B. Bouyssiere, H. Zhou and R. G. Bryant, Multi-dimensional Nuclear Magnetic Resonance characterizations of dynamics and saturations of brine/crude oil/mud filtrate mixtures confined in rocks: The role of asphaltene, Energy Fuels 28 (2014), 1629-1640.

[19] Z. Yang and G. J. Hirasaki, NMR measurement of bitumen at different temperatures, J. Magn. Reson. 192 (2008), 280-293.

[20] H. Cramér, Mathematical Methods of Statistics, Princeton University Press, 1963.

[21] J. A. Jones, P. Hodgkinson, A. L. Barker and P. J. Hore, Optimal sampling strategies for the measurement of spin-spin relaxation times, J. Magn. Reson. Series B 113 (1996), 25-34.

[22] Y. Zhang, H. N. Yeung, M. O'Donnell and P. L. Carson, Determination of sample time for T_1 measurement, J. Magn. Reson. Imaging 8 (1998), 675-681.

[23] P. Hodgkinson, K. J. Holmes and P. J. Hore, Selective data acquisition in NMR. The quantification of anti-phase scalar couplings, J. Magn. Reson. Series A 18 (1996), 18-30.

[24] D. C. Alexander, A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features, Magn. Reson. Med. 60 (2008), 439-448.

[25] P. Becher and M. N. Yudenfreund (Ed.), Emulsions, Lattices and Dispersions, Marcel Dekker, 1978.

[26] K. J. Packer and C. Rees, Pulsed NMR studies of restricted diffusion. I. Droplet size distributions in emulsions, J. Colloid Interface Sci. 40 (1972), 206-218.

[27] P. T. Callaghan, K. W. Jolley and R. S. Humphrey, Diffusion of fat and water in cheese as studied by pulsed field gradient Nuclear Magnetic Resonance, J. Colloid Interface Sci. 93 (1983), 521-529.

[28] J. C. van Den Enden, D. Waddington, H. van Aalst, C. G. Van Kralingen and K. J. Packer, Rapid determination of water droplet size distributions by PFG-NMR, J. Colloid Interface Sci. 140 (1990), 105-113.

[29] K. G. Hollingsworth and M. L. Johns, Measurement of emulsion droplet sizes using PFG NMR and regularization methods, J. Colloid Interface Sci. 258 (2003), 383-389.

[30] I. Fourel, J. P. Guillement and D. Le Botlan, Determination of water droplet size distributions by low resolution PFG-NMR, J. Colloid Interface Sci. 164 (1994), 48-53.

[31] G. J. W. Goudappel, J. P. M. van Duynhoven and M. M. Mooren, Measurement of oil droplet size distributions in food oil/water emulsions by time domain pulsed field gradient NMR, J. Colloid Interface Sci. 239 (2001), 535-542.

[32] M. L. Johns and L. F. Gladden, Sizing of emulsion droplets under flow using flow-compensating NMR-PFG techniques, J. Magn. Reson 154 (2002), 142-145.

[33] E. S. Rajagopal, Statistical theory of particle size distributions in emulsions and suspensions, Kolloid Z. 162 (1959), 85-92.

[34] J. C. Thomas, The determination of log normal particle size distributions by dynamic light scattering, J. Colloid Interface Sci. 117 (1987), 187-192.

[35] G. van Dalen, Determination of the water droplet size distribution of fat spreads using confocal scanning laser microscopy, J. Microsc. 208 (2002), 116-133.

[36] Y. Katz and U. Nevo, Quantification of pore size distribution using diffusion NMR: Experimental design and physical insights, J. Chem. Phys. 140 (2014), 164201.

[37] M. Nilsson, M. A. Connell, A. L. Davis and G. A. Morris, Biexponential fitting of diffusion-ordered NMR data: Practicalities and limitations, Anal. Chem. 78 (2006), 3040-3045.

[38] M. S. Shazeeb, C. H. Sotak, Limitations in biexponential fitting of NMR inversion-recovery curves, J. Magn. Reson. 276 (2017), 14-21.

[39] R. V. Mulkern, M. Balasubramanian and S. E. Maier, On the perils of multiexponential fitting of diffusion MR data, J. Magn. Reson. Imaging 45 (2016), 1545-1547.

[40] G. Pagès, A. Bonny, V. Gilard and M. Malet-Martino, Pulsed field gradient NMR with sigmoid shape gradient sampling to produce more detailed diffusion ordered spectroscopy maps of real complex mixtures: Examples with medicine analysis, Anal. Chem. 88 (2016), 3304-3309.

[41] R. Bai, D. Benjamini, J. Cheng and P. J. Basser, Fast, accurate 2D-MR relaxation exchange spectroscopy (REXSY): Beyond compressed sensing, J. Chem. Phys. 145 (2016), 154202.

[42] A. Reci, Signal processing in magnetic resonance applications, PhD thesis, 2018, *to be submitted*.

[43] E. O. Stejskal and J. E. Tanner, Spin diffusion measurements: Spin echoes in the presence of time-dependent field gradient, J. Chem. Phys. 42 (1965), 288-292.

[44] J. S. Murday and R. M. Cotts, Self-diffusion coefficient of liquid lithium, J. Chem. Phys. 48 (1968), 4938-4945.

[45] M. Alderliesten, Mean particle diameters. Part I: Evaluation of definition systems, Part. Part. Syst. Charact. 7 (1990), 233-241.

[46] T. Hatch and S. P. Choate, Statistical description of the size properties of non-uniform particulate substances, J. Franklin Inst. 207 (1929), 369-387.

[47] R. A. Fisher, The Design of Experiments, Oliver and Boyd, 1953.

[48] J. P. Norton, An Introduction to Identification, Academic Press, 1986.

[49] G. Strang, Introduction to Linear Algebra, Wellesley-Cambridge Press, 4th ed., 2009.

[50] O. V. Petrov and S. Stapf, Parametrization of NMR relaxation curves in terms of logarithmic moments of the relaxation time distribution, J. Magn. Reson. 279 (2017), 29-38.

[51] J. E. Tanner, Use of the stimulated echo in NMR diffusion studies, J. Chem. Phys. 52 (1970), 2523-2526.

[52] D. E. O'Reilly and E. M. Peterson, Self-diffusion coefficients and rotational correlation times in polar liquids. III. Toluene, J. Chem. Phys. 56 (1972), 2262-2266.

[53] G. H. Weiss, R. K. Gupta, J. A. Ferretti and E. D. Becker, The choice of optimal parameters for measurement of spin-lattice relaxation times. I. Mathematical formulation, J. Magn. Reson. 37 (1980), 369-379.

[54] B. Efron, The Jackknife, the Bootstrap and Other Resampling Plans, SIAM, 1982.

[55] P. T. Callaghan, Principles of Nuclear Magnetic Resonance Microscopy, Oxford University Press, 1993.

[56] J. R. Taylor, An Introduction to Error Analysis, University Science Books, 1997.

Figure captions

Fig. 1. (a) Experimental data for the attenuation of the aromatic peak of toluene (chemical shift $\delta \sim 7.1$ ppm with respect to TMS, shown in the embedded figure) from the PFG NMR diffusion experiment of an emulsion of toluene in water. The numerical fit is estimated from the best fit lognormal distribution for the emulsion droplet size distribution. The estimated noise level from the fit is $\sigma = 5.6 \times 10^{-4}$. (b) Best fit lognormal distribution for the emulsion droplet size distribution with $\bar{a}_{0,0} = 1.09$, $s_g = 0.38$ ($\bar{a}_{1,0} = 3.2\ \mu$m, $s = 1.2\ \mu$m).

Fig. 2. (a) Contour map of the variation of the objective function, χ, with respect to the parameters of the family of linear sampling patterns defined in Eq. (12) with $n = 32$, as predicted by the CRLB theory. (b) Contour map of the variation of the objective function, χ, with respect to the parameters of the class of linear sampling patterns defined in Eq. (12) with $n = 32$, obtained from the experimental data shown in Fig. 1(a). The best linear sampling pattern in (a) is obtained for $r = 0.67$, $g_n = 159$ G cm^{-1}, and is identified as +, at which point $\chi = 1.03 \times 10^{-3}$.

Fig. 3. (a) Contour map of the variation of the objective function, χ, with respect to the parameters of the family of logarithmic sampling patterns defined in Eq. (13) with $n = 32$, as predicted by the CRLB theory. (b) Contour map of the variation of the objective function, χ, with respect to the parameters of the class of logarithmic sampling patterns defined in Eq. (13) with $n = 32$, obtained from the experimental data shown in Fig. 1(a). The best logarithmic sampling pattern in (a) is obtained for $r = 0.23$, $g_n = 161$ G cm^{-1}, and is identified as +, at which point $\chi = 1.03 \times 10^{-3}$.

Fig. 4. (a) Variation of the objective function, χ, with the noise level, σ_n, as calculated by Monte Carlo type simulations and as predicted by the CRLB theory. The sampling pattern used was the best linear sampling pattern for $n = 32$ with $r = 0.67$, $g_n = 159$ G cm^{-1}. (b) Contour map showing the percentage increase in the objective function, $\Delta\chi$, caused by the mis-design of the linear sampling pattern defined in Eq. (12) because of a percentage error in guessing the arithmetic mean, $\Delta\bar{a}_{1,0}$ (μm), and standard deviation, Δs (μm), of the emulsion droplet size distribution. The mis-designed sampling pattern used was the best

linear obtained in Fig. 2(a) for $n = 32$ with $r = 0.67$, $g_n = 159$ G cm^{-1} and the noise level used in the CRLB predictions was $\sigma = 5.6 \times 10^{-4}$.

Fig. 5. (a) Variation of the minimum objective function, χ, with the number of data points in a linear sampling pattern, n, calculated by the CRLB theory for the case of a fixed acquisition time experiment (◊). Variation of the reduction in acquisition time, t / t_{def}, with n calculated by the CRLB theory for the case of a fixed χ experiment (×). The default acquisition time, t_{def}, was calculated for a default sampling pattern of 16 equidistant points and $g_n = 300$ G cm^{-1}. The lines are included to guide the eye. (b) Best linear sampling pattern for $n = 7$ with $r = 0.98$, $g_n = 147$ G cm^{-1}, which according to (a) is the optimal sampling pattern for both a fixed acquisition time experiment and a fixed χ experiment. The line is included to guide the eye.

Figure 1

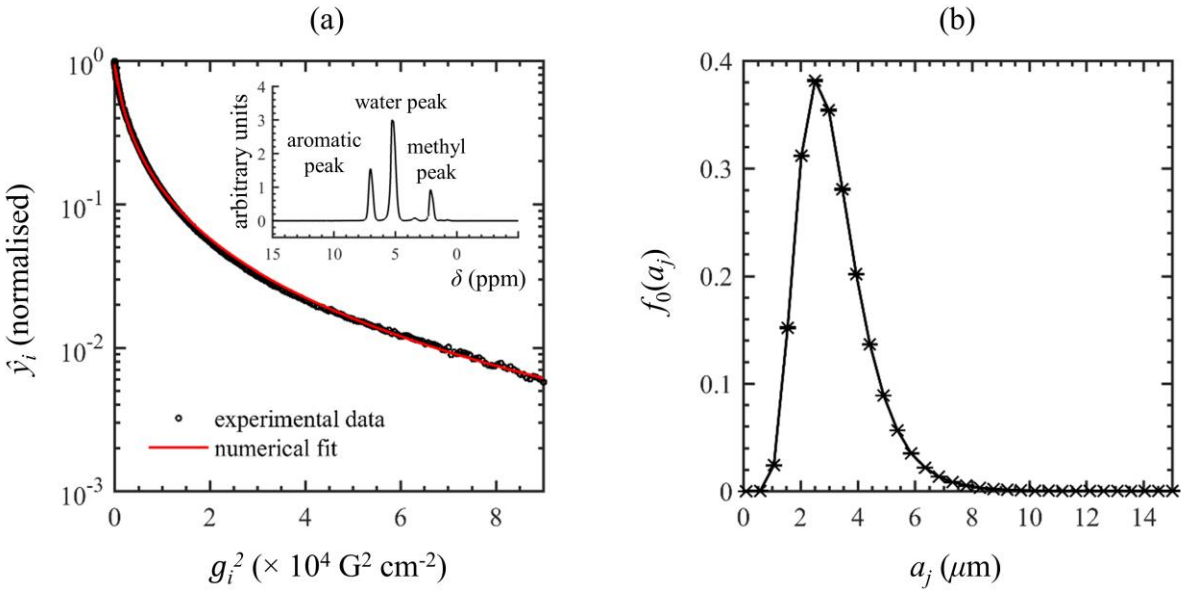

Figure 2

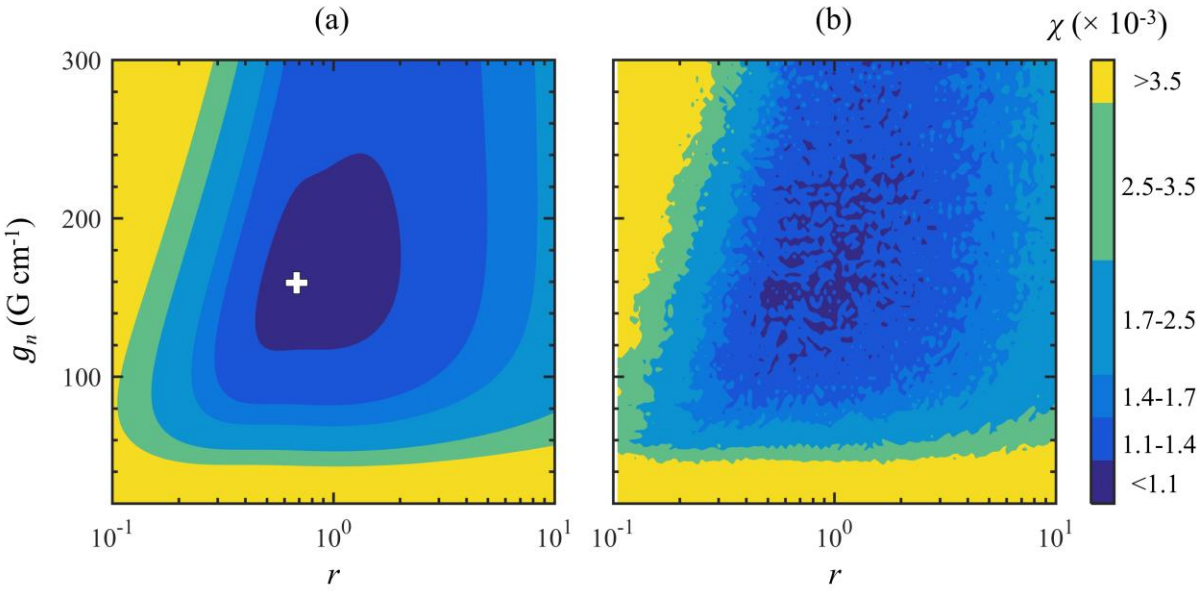

Figure 3

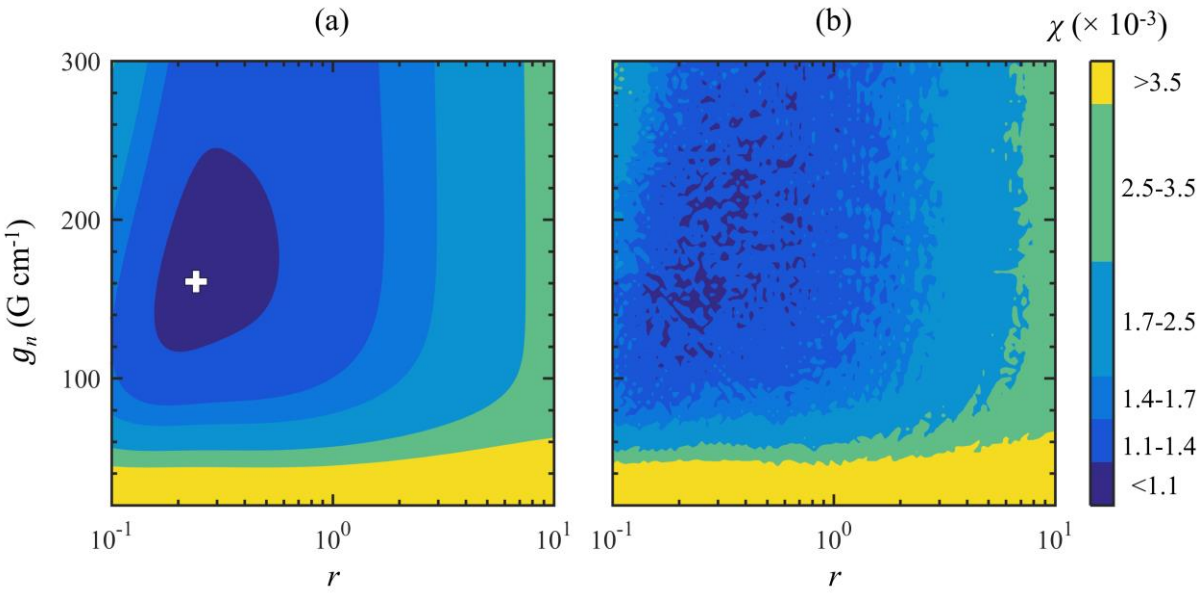

Figure 4

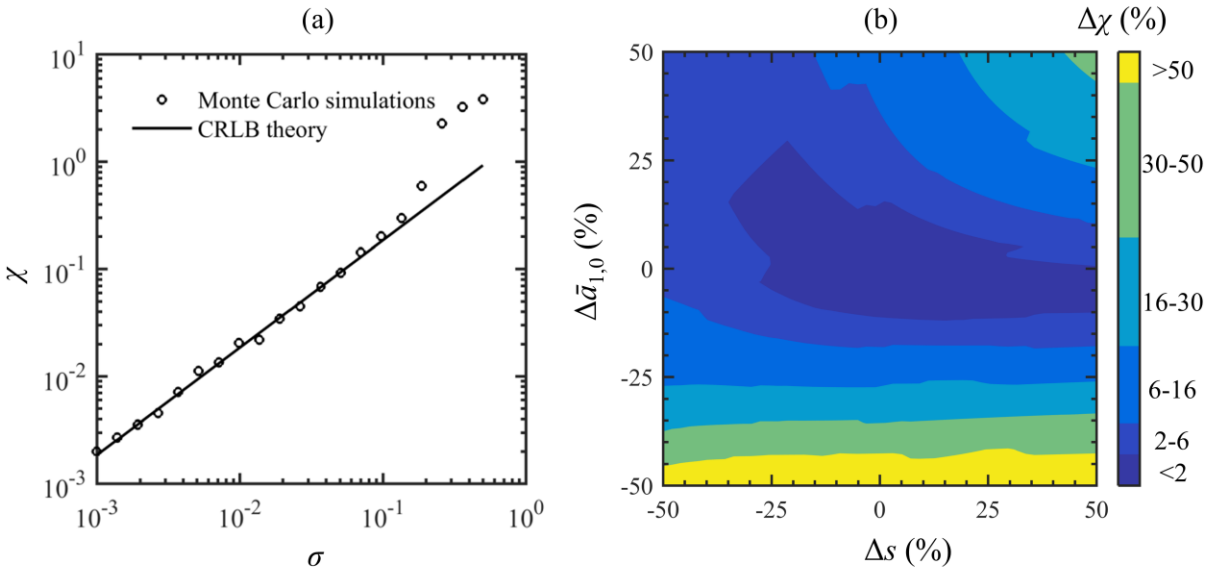

Figure 5